Eco 5316 Time Series Econometrics

Lecture 21 Vector Error Correction (VEC) Models

Motivation

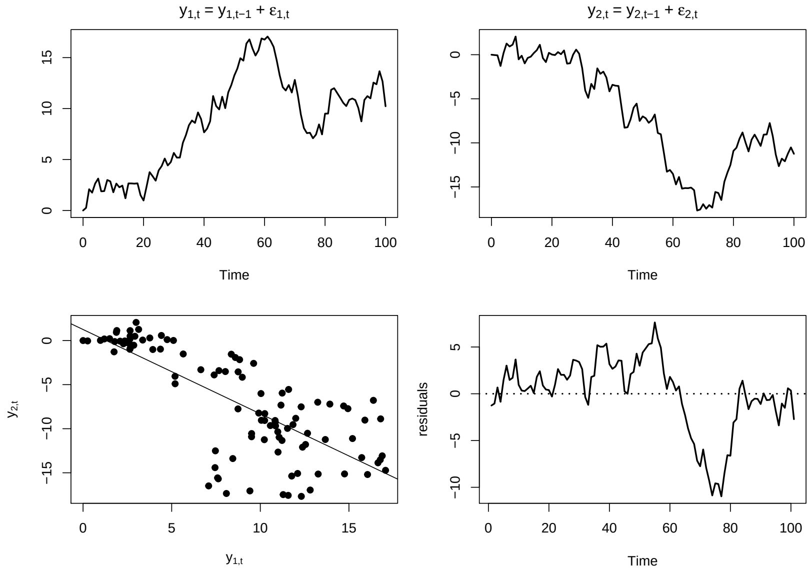
- in a VAR model all variables need to be weakly stationary, and we can estimate it equation by equation using standard OLS
- we will need different methodology for nonstationary time series, since spurious regression problem can arise with standard OLS when the times series are nonstationary

Spurious Regression

- **spurious regression** problem - running and OLS with integrated variables can yield significant coefficients, even though the variables are not related
- example: suppose that $y_{i,t} = y_{i,t-1} + \varepsilon_{i,t}$ for $i = 1, 2$ and that we estimate a simple OLS $y_{2,t} = \beta_0 + \beta_1 y_{1,t} + e_t$
- if $T \to \infty$ then $\beta_1 \not\to 0$, and in addition t-statistics $\to \pm\infty$ and $R^2 \to 1$
- residuals from the OLS will show significant serial correlation

- bottom line: with nonstationary time series we have to be vary careful with OLS regressions, correlation does not necessarily mean causality http://tylervigen.com/spurious-correlations

Example: Spurious Regression

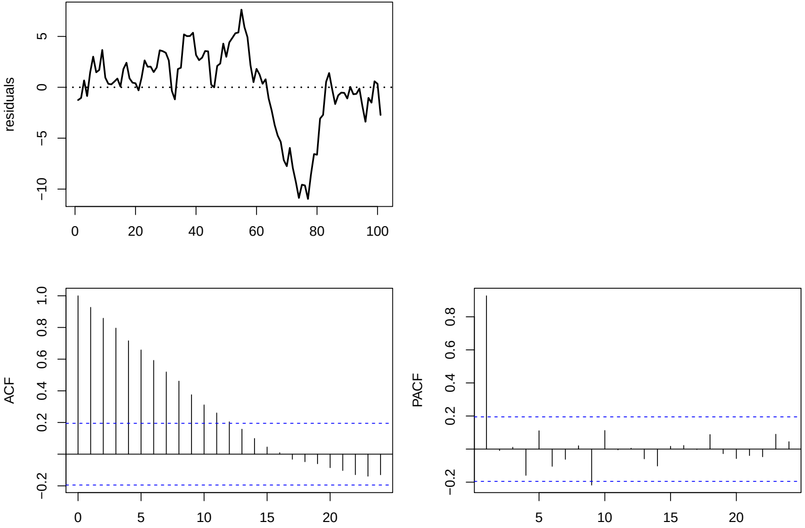


Example: Spurious Regression

```r
myOLS <- lm(y2 ~ y1)
summary(myOLS)
```

```
##
## Call:
## lm(formula = y2 ~ y1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9681  -1.1017   0.5366   2.4119   7.6268
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.25535    0.82054    1.53    0.129
## y1          -0.95641    0.08371  -11.43   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.952 on 99 degrees of freedom
## Multiple R-squared:  0.5687, Adjusted R-squared:  0.5643
## F-statistic: 130.5 on 1 and 99 DF,  p-value: < 2.2e-16
```

Example: Spurious Regression



Cointegration and Error Correction Models

- economic models often feature
  (1) economic variables related to each other through a long-run equilibrium relationship
  (2) forces that push the variables toward this equilibrium if there is a temporary deviation from it (for long-run equilibrium to exist, movements of some of the variables must respond to the magnitude and direction of the deviation from it)
- this motivates the concepts of cointegration and error correction

Cointegration

some examples of relationships predicted by theory

- real wages and labor productivity $\frac{w}{p} = (1-\alpha)\frac{Y}{H}$
- money demand: stock of money, price level, real GDP, nominal interest rate $\frac{M}{p} = L(Y, i)$
- purchasing power parity: prices of the same good in two countries and the exchange rate $p_t^* = e_t p_t$
- prices of same stock traded on two stock exchanges
- short run and long run interest rates
- permanent income and consumption expenditures $C_t = \gamma Y_t^p$

testing for cointegration is then essentially investigating whether particular theory is consistent with data

Cointegration

- recall: a variable y_t is $I(d)$, **integrated of order** d, if it is nonstationary and differencing it d times produces a stationary variable $\Delta^d y_t$
- vector of variables $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{nt})$ is said to be **cointegrated of order** d, b, denoted by $CI(d, b)$ if
  1. all components are integrated of order d
  2. there exists a cointegrating vector $\beta = (\beta_1, \ldots, \beta_n)' \neq 0$ such that $\beta' \boldsymbol{y}_t = \beta_1 y_{1t} + \ldots + \beta_n y_{nt}$ is integrated of order $d - b$
- cointegrating vector is not unique, $\lambda\beta$ also satisfies the condition for any $\lambda > 0$; we usually normalize $\beta_1 = 1$
- number of cointegrating vectors is called the **cointegrating rank** of $\boldsymbol{y}_t$
- for n variables there can be up to $n-1$ linearly independent cointegrating vectors
- in economics most nonstationary time series are $I(1)$ and so $C(1, 1)$ is the most common case of cointegration

Cointegration

- real wage and labor productivity example: if

$$\frac{w}{p} = (1-\alpha)\frac{Y}{H}$$

  then

$$\log\frac{w}{p} = \log(1-\alpha) + \log\frac{Y}{H}$$

- in the data this will not hold all the time so

$$\log\frac{w}{p} = \log(1-\alpha) + \log\frac{Y}{H} + e_t$$

  but theory suggests that $\log\frac{w}{p}$ and $\log\frac{Y}{H}$ should be $I(1)$ due to technological progress and e_t should be $I(0)$ weakly stationary

- theory thus suggests that if $\boldsymbol{y}_t = (\log\frac{w_t}{p_t}, \log\frac{Y_t}{H_t})'$ then

$$e_t = \log\frac{w}{p} - \log\frac{Y}{H} - \log(1-\alpha)$$

  is $I(0)$ so the cointegrating vector is $\boldsymbol{\beta} = (\beta_1, \beta_2) = (1, -1)$ and we should include a constant when testing for cointegration

Cointegration

- driving force behind cointegration - variables share a common stochastic trend; e.g. real wages and labor productivity both grow because of technological progress that affects both of them
- consider the case with two $I(1)$ time series $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$ where

$$y_{i,t} = \delta_i + \mu_{i,t} + x_{i,t} \qquad \text{for } i = 1,2$$

  where $\mu_{i,t} = \sum_{j=0}^{t} \varepsilon_{i,t}$ are the stochastic trend components and $x_{i,t}$ are some weakly stationary $I(0)$ processes
- these two are cointegrated if there exists vector $\boldsymbol{\beta} = (\beta_1, \beta_2)'$ such that

$$\beta_1 y_{1,t} + \beta_2 y_{2,t}$$

  is $I(0)$, i.e. weakly stationary
- we have

$$\begin{aligned}&\beta_1 y_{1,t} + \beta_2 y_{2,t}\\ &= \left(\beta_1\delta_1 + \beta_2\delta_2\right) + \left(\beta_1\mu_{1,t} + \beta_2\mu_{2,t}\right) + \left(\beta_1 x_{1,t} + \beta_2 x_{2,t}\right)\end{aligned}$$

  which is only stationary if $\beta_1\mu_{1,t} + \beta_2\mu_{2,t} = 0$ so that $\mu_{1,t} = -\beta_2/\beta_1\mu_{2,t}$
- thus to be $C(1,1)$ cointegrated, $y_{1,t}$ and $y_{2,t}$ must share the same stochastic trend, and cointegrating vector $\boldsymbol{\beta}$ removes this stochastic trend from the linear combination of $y_{1,t}$ and $y_{2,t}$

Cointegration Test - Engle-Granger Methodology

- main idea for Engle-Granger test for cointegration: test whether residuals from an OLS contain a unit root, if they do, there's no cointegration, just spurious regression
- example: consider two variables $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$
  - step 1: test whether variables $y_{1,t}$ and $y_{2,t}$ are $I(1)$
  - step 2: estimate one of the models

$$y_{1,t} = \beta_2 y_{2,t} + e_t$$
$$y_{1,t} = \delta_0 + \beta_2 y_{2,t} + e_t$$
$$y_{1,t} = \delta_0 + \delta_1 t + \beta_2 y_{2,t} + e_t$$

  - step 3: test residuals e_t for the presence unit root
    - if we can not reject H_0 of unit root in residuals, we can not reject the H_0 that $y_{1,t}$ and $y_{2,t}$ are not cointegrated
    - rejecting H_0 of unit root in residuals means rejecting that $y_{1,t}$ and $y_{2,t}$ are not cointegrated

Cointegration Test - Engle-Granger Methodology

- Engle-Granger Methodology has several significant drawbacks
  - usual critical values can not be applied when testing for a unit root in residuals, because coefficients $\beta_2, \ldots, \beta_n$ are unknown and were estimated
  - critical values depend on deterministic terms used and number of variables
  - exchanging $y_{1,t}$ and $y_{2,t}$ in the OLS may lead to contradictory results
  - no way to test for cointegrating rank
  - no easy way to test various restrictions on coefficients $\beta_2, \ldots, \beta_n$
- because of these drawbacks Johansen's methodology is generally preferred to Engle-Granger Methodology

Cointegration Test - Johansen's Methodology

preview of main steps involved in Johansen's Methodology

- specify and estimate a VAR(p) model for $\boldsymbol{y}_t$ (in levels, not in differences)
- determine number of cointegrating vectors using trace and max eigenvalue tests
- estimate a vector error correction model by maximum likelihood

Vector Error Correction Model

- in an error-correction model (ECM), short-term dynamics of variables in the system is influenced by the size of the deviation from long-run equilibrium
- suppose that two $I(1)$ variables $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$ are $CI(1,1)$ cointegrated with cointegrating vector $\boldsymbol{\beta} = (1, \beta_2)'$ so that $y_{1,t} + \beta_2 y_{2,t}$ is $I(0)$
- consider a VAR(1) model $\boldsymbol{y}_t = \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t$ or

$$
\begin{aligned}
y_{1,t} &= a_{11} y_{1,t-1} + a_{12} y_{2,t-1} + \varepsilon_{1,t} \\
y_{2,t} &= a_{21} y_{1,t-1} + a_{22} y_{2,t-1} + \varepsilon_{2,t}
\end{aligned}
$$

Vector Error Correction Model

- subtract $y_{i,t-1}$ from equation i to get $\Delta \boldsymbol{y}_t = \boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \varepsilon_t$ where $\boldsymbol{\Pi} = -(\boldsymbol{I} - \boldsymbol{A}_1)$ or equivalently

$$\begin{aligned}
\Delta y_{1,t} &= -(1-a_{11})\left(y_{1,t-1} - a_{12}/(1-a_{11}) y_{2,t-1}\right) + \varepsilon_{1,t} \\
\Delta y_{2,t} &= a_{21}\left(y_{1,t-1} - (1-a_{22})/a_{21} y_{2,t-1}\right) + \varepsilon_{2,t}
\end{aligned}$$

- the LHS variables $\Delta y_{1,t}$ and $\Delta y_{2,t}$ are $I(0)$, the RHS are $I(0)$ only if $(1-a_{11}) = 0$, $a_{21} = 0$, or if $\beta_2 = -a_{12}/(1-a_{11}) = -(1-a_{22})/a_{21}$
- we have obtained a simple vector error correction (VEC) model

$$\begin{aligned}
\Delta y_{1,t} &= \alpha_1 (y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{1,t} \\
\Delta y_{2,t} &= \alpha_2 (y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{2,t}
\end{aligned}$$

where $\alpha_1 = -(1-a_{11})$, $\alpha_2 = a_{21}$ and $\beta_2 = -a_{12}/(1-a_{11}) = -(1-a_{22})/a_{21}$

Vector Error Correction Model

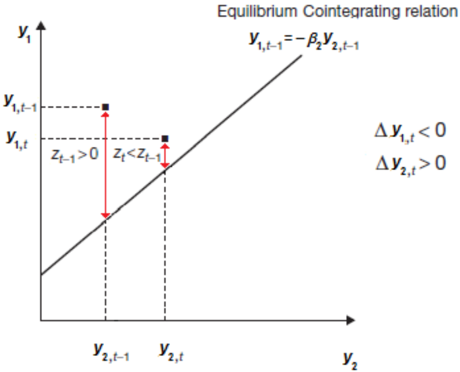
- consider the simple vector error correction model we obtained

$$
\begin{aligned}
\Delta y_{1,t} &= \alpha_1 (y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{1,t} \\
\Delta y_{2,t} &= \alpha_2 (y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{2,t}
\end{aligned}
$$

- $(y_{1,t-1} + \beta_2 y_{2,t-1})$ is referred to as the **error correction term**
- **adjustment parameters** α_1, α_2 determine the speed of return to long run equilibrium, the larger they are in absolute value, the less persistent deviations from long-run equilibrium become
- for long run relationship to be stable $\alpha_1 \leq 0$, $\alpha_2 \geq 0$ needs to be satisfied, and at least one of them can not be equal 0
- if for example $y_{1,t-1} + \beta_2 y_{2,t-1} > 0$, then $y_{1,t-1}$ is too high and $y_{2,t-1}$ too low compared to the long run equilibrium, and if $\alpha_1 > 0$ and $\alpha_2 < 0$ then y_1 would be growing and y_2 would be declining, taking the system even further away from the long run equilibrium

Vector Error Correction Model

- consider two processes $y_{1,t}$ and $y_{2,t}$, with cointegrating relation $y_{1,t} + \beta_2 y_{2,t}$
- suppose that at time $t-1$ the the system is out of equilibrium with $z_{t-1} = y_{1,t-1} + \beta_2 y_{2,t-1} > 0$



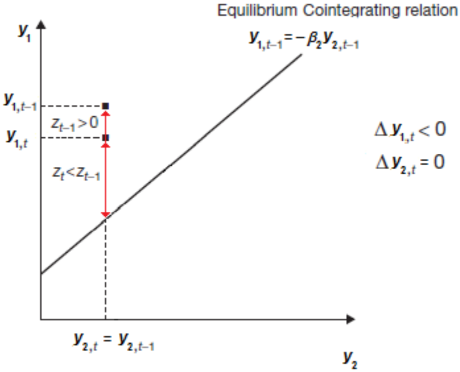
Vector Error Correction Model

- consider two processes $y_{1,t}$ and $y_{2,t}$, with cointegrating relation $y_{1,t}+\beta_2 y_{2,t}$
- suppose that at time $t-1$ the the system is out of equilibrium with $z_{t-1} = y_{1,t-1}+\beta_2 y_{2,t-1} > 0$
- cointegrating relation exercises a "gravitational pull": in period t system will partially self-correct the disequilibrium of period $t-1$, and over time gradually move toward the equilibrium
- to reach $(y_{1,t}, y_{2,t})$ from $(y_{1,t-1}, y_{2,t-1})$, y_1 has decreased and y_2 has increased, so $\Delta y_{1,t} < 0$ and $\Delta y_{2,t} > 0$
- note that in period t there is still a disequilibrium z_t, but of smaller magnitude, $|z_t| < |z_{t-1}|$
- if there are no other shocks in the following periods, the system will keep correcting the disequilibrium error until it reaches the equilibrium path, and once there, it will not move out

- it is possible for one of the adjustment parameters to be zero: if $\alpha_1 < 0, \alpha_2 = 0$ then $y_{2,t}$ is a pure random walk and all the adjustment occurs in $y_{1,t}$; in this case $y_{2,t}$ is said to be **weakly exogenous**

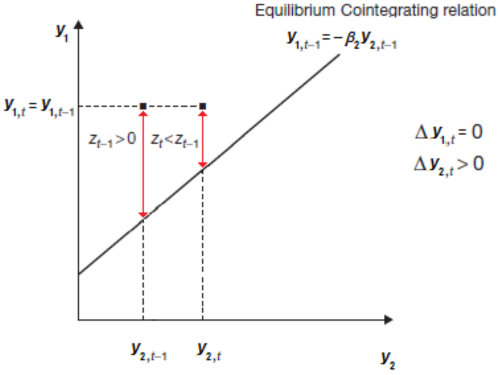
Vector Error Correction Model

- if $\alpha_1 < 0$ and $\alpha_2 = 0$ adjustment only takes place in y_1, while y_2 remains the unchanged
- e.g. if y_2 is income and y_1 consumption expenditures, this would mean that consumption drops over time if it is unsustainably high, and income remains same over time



Vector Error Correction Model

- if $\alpha_1 = 0$ and $\alpha_2 > 0$ adjustment only takes place in y_2, while y_1 remains the unchanged
- e.g. if y_2 is production and y_1 consumption, this would mean that if consumption is too high, it will remain unchanged, but production will grow over time



Vector Error Correction Model

- consider the simple vector error correction model we obtained

$$\Delta y_{1,t} = \alpha_1(y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{1,t}$$
$$\Delta y_{2,t} = \alpha_2(y_{1,t-1} + \beta_2 y_{2,t-1}) + \varepsilon_{2,t}$$

- more compactly we can write

$$\Delta \boldsymbol{y}_t = \mathbf{\Pi} \boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t$$

where

$$\mathbf{\Pi} = \begin{pmatrix} \alpha_1 & \alpha_1\beta_2 \\ \alpha_2 & \alpha_2\beta_2 \end{pmatrix}$$

so that $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)'$ and $\boldsymbol{\beta} = (1, \beta_2)'$

Vector Error Correction Model

- consider now $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{n,t})'$ that are $I(1)$ and follow VAR(p)
$$\boldsymbol{y}_t = \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{A}_2 \boldsymbol{y}_{t-2} + \ldots + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \varepsilon_t$$
- first add and subtract $\boldsymbol{A}_p \boldsymbol{y}_{t-p+1}$ to get
$$\boldsymbol{y}_t = \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{A}_2 \boldsymbol{y}_{t-2} + \ldots + \boldsymbol{A}_{p-2} \boldsymbol{y}_{t-p+2} + (\boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \boldsymbol{y}_{t-p+1} - \boldsymbol{A}_p \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$
- next add and subtract $(\boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \boldsymbol{y}_{t-p+2}$ to get
$$\begin{aligned} \boldsymbol{y}_t = {} & \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + \boldsymbol{A}_2 \boldsymbol{y}_{t-2} + \ldots \\ & + (\boldsymbol{A}_{p-2} + \boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \boldsymbol{y}_{t-p+2} - (\boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \Delta \boldsymbol{y}_{t-p+2} - \boldsymbol{A}_p \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t \end{aligned}$$
- by adding and subtracting $(\boldsymbol{A}_{p-i+1} + \ldots + \boldsymbol{A}_p) \boldsymbol{y}_{t-p+i}$ for $i = 1, \ldots, p-1$
$$\begin{aligned} \boldsymbol{y}_t = {} & (\boldsymbol{A}_1 + \ldots + \boldsymbol{A}_p) \boldsymbol{y}_{t-1} \\ & - (\boldsymbol{A}_2 + \ldots + \boldsymbol{A}_p) \Delta \boldsymbol{y}_{t-1} - \ldots - (\boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \Delta \boldsymbol{y}_{t-p+2} - \boldsymbol{A}_p \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t \end{aligned}$$
- finally subtract $\boldsymbol{y}_{t-1}$ to get
$$\begin{aligned} \Delta \boldsymbol{y}_t = {} & -(\boldsymbol{I} - \boldsymbol{A}_1 - \ldots - \boldsymbol{A}_p) \boldsymbol{y}_{t-1} \\ & - (\boldsymbol{A}_2 + \ldots + \boldsymbol{A}_p) \Delta \boldsymbol{y}_{t-1} - \ldots - (\boldsymbol{A}_{p-1} + \boldsymbol{A}_p) \Delta \boldsymbol{y}_{t-p+2} - \boldsymbol{A}_p \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t \end{aligned}$$
- more compactly: $\Delta \boldsymbol{y}_t = \boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$ where $\boldsymbol{\Pi} = -(\boldsymbol{I} - \boldsymbol{A}_1 - \ldots - \boldsymbol{A}_p)$ and $\boldsymbol{\Gamma}_i = -(\boldsymbol{A}_{i+1} + \ldots + \boldsymbol{A}_p)$

Vector Error Correction Model

- if $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{n,t})'$ is a vector of $I(1)$ variables its VEC representation is

$$\Delta \boldsymbol{y}_t = \boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

  where $\boldsymbol{\Pi} = \alpha \beta'$ and in addition also $\boldsymbol{\Pi} = -(\boldsymbol{I} - \boldsymbol{A}_1 - \ldots - \boldsymbol{A}_p)$ and $\boldsymbol{\Gamma}_i = -(\boldsymbol{A}_{i+1} + \ldots + \boldsymbol{A}_p)$ for $i = 1, \ldots, p-1$
- **Granger representation theorem**: for any set of $I(1)$ variables error correction representation exists if and only if they are cointegrated

Vector Error Correction Model

- consider a VEC model, augmented by a deterministic term $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$

$$\Delta \boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \boldsymbol{\varepsilon}_t$$

  where $\boldsymbol{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$ and in addition also $\boldsymbol{\Pi} = -(\boldsymbol{I} - \boldsymbol{A}_1 - \ldots - \boldsymbol{A}_p)$ and $\boldsymbol{\Gamma}_i = -(\boldsymbol{A}_{i+1} + \ldots + \boldsymbol{A}_p)$ for $i = 1, \ldots, p-1$

- with n variables there are up to $n-1$ cointegrating vectors, so $\boldsymbol{\beta}$ is in general a matrix with n columns and number of rows equal to number of cointegrating vectors (i.e. number of long-run relationships)
- similarly $\boldsymbol{\alpha}$ is in general a matrix with n rows and number of columns equal to number of cointegrating vectors, element in row i column j represents the correction of variable i to a deviation in j long-run relationship
- note that if $\boldsymbol{\Pi} = \boldsymbol{0}$ VEC model above become a reduced form VAR(p) model estimated on differenced data
- if $\boldsymbol{\Pi}$ contains non-zero elements estimating a VAR on differenced data $\Delta \boldsymbol{y}_t$ leads to omitted variable bias - it is not appropriate to estimate a VAR using first differences if the variables are cointegrated

Cointegration Test - Johansen's Methodology

- recall: rank of a matrix is defined as the number of linearly independent rows it contains
- since $\mathbf{\Pi}$ has as many linearly independent rows as there are cointegrating vectors β, it is possible to test for cointegration using rank of matrix $\mathbf{\Pi}$
  - if $\text{rank}(\mathbf{\Pi}) = 0$ then $\boldsymbol{y}_t$ are $I(1)$ but not cointegrated
  - if $0 < \text{rank}(\mathbf{\Pi}) < n$ then $\boldsymbol{y}_t$ are cointegrated with $r = \text{rank}(\mathbf{\Pi})$ linearly independent long-run relationships
  - if $\text{rank}(\mathbf{\Pi}) = n$ then $\boldsymbol{y}_t$ must actually be $I(0)$ weakly stationary and there is no cointegration among them
- e.g. in the previous bivariate example where we had

$$\mathbf{\Pi} = \begin{pmatrix} \alpha_1 & \alpha_1\beta_2 \\ \alpha_2 & \alpha_2\beta_2 \end{pmatrix}$$

the two rows are linearly dependent, $\text{rank}(\mathbf{\Pi}) = 1$, since there is only one cointegrating vector $\boldsymbol{\beta} = (1, \beta_2)'$

Cointegration Test - Johansen's Methodology

- let $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots \hat{\lambda}_n$ be the estimated eigenvalues of $\mathbf{\Pi}$
- if $\text{rank}(\mathbf{\Pi}) = r$ then $\hat{\lambda}_{r+1}, \dots, \hat{\lambda}_n$ should be small, close to 0
- to test H_0: $\text{rank}(\mathbf{\Pi}) = r$ against H_A: $\text{rank}(\mathbf{\Pi}) > r$ for $r = 0, 1, \dots, n-1$ we use **trace statistic**

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \log(1 - \hat{\lambda}_i)$$

- to test H_0: $\text{rank}(\mathbf{\Pi}) = r$ against H_A: $\text{rank}(\mathbf{\Pi}) = r+1$ for $r = 0, 1, \dots, n-1$ we use **maximum eigenvalue statistic**

$$\lambda_{max}(r) = -T \log(1 - \hat{\lambda}_{r+1})$$

- results of trace and max eigenvalue test may be contradictory; if that happens max eigenvalue test is usually prioritized

Cointegration Test - Johansen's Methodology

- for each of the two tests we follow a sequential procedure
- for example with trace test we would proceed as follows
  - step 1: test H_0 : rank($\mathbf{\Pi}$) $= 0$ against H_A : rank($\mathbf{\Pi}$) > 0, if H_0 is not rejected we conclude that $\boldsymbol{y}_t$ are not cointegrated, otherwise we move to the next step
  - step 2: test H_0 : rank($\mathbf{\Pi}$) $= 1$ against H_A : rank($\mathbf{\Pi}$) > 1, if H_0 is not rejected we conclude that there is one cointegrating vector, otherwise we move to next step
  - in general in step i for $i = 1, 2, \ldots, n-1$ we test H_0 : rank($\mathbf{\Pi}$) $= i$ against H_A : rank($\mathbf{\Pi}$) $> i$, if H_0 is not rejected we conclude that there are i cointegrating vectors, otherwise we move to next step
  - this procedure is continued until the null is not rejected
- with max eigenvalue test the H_A is different, but the overall sequential approach is the same

Cointegration Test - Johansen's Methodology, An Example

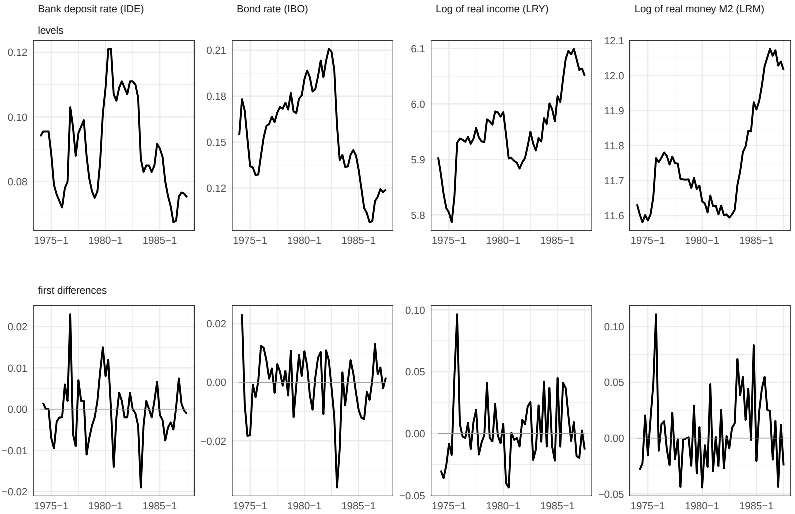
- data for Denmark, 1974Q1-1987Q3, $\boldsymbol{y}_t = (\log(M2_t/P_t), \log Y_t, i_t^b, i_t^d)'$ where $\log(M2_t/P_t)$ is log of money supply M2 deflated by price index, $\log Y_t$ is log of real income, i_t^b is bond rate, i_t^d is deposit rate
- based on unit root tests all series appear to be $I(1)$

Cointegration Test - Johansen's Methodology, An Example

```
library(magrittr)
library(tidyverse)
library(timetk)
library(zoo)
library(lubridate)
library(urca)
library(vars)
library(ggfortify)
library(egg)
library(qqplotr)

# set default theme for ggplot2
theme_set(theme_bw() +
          theme(strip.text.x = element_text(hjust = 0),
                strip.text.y = element_text(hjust = 1),
                axis.ticks = element_blank(),
                strip.background = element_blank()))

# load data to estimate money demand function of Denmark, 1974Q1 to 1987Q3
data(denmark)

# convert data into tibble format
denmark_tbl <-
    denmark %>%
    as_tibble() %>%
    mutate(yearq = as.yearqtr(ENTRY, format = "%Y:%q")) %>%
    dplyr::select(yearq, LRM, LRY, IBO, IDE)

# convert data into ts format
denmark_ts <-
    denmark_tbl %>%
    tk_ts(select = -yearq, start = year(.$yearq[1]), frequency = 4)
```

Cointegration Test - Johansen's Methodology, An Example

Data set for Denmark from Johansen & Juselius (1990)



Cointegration Test - Johansen's Methodology, An Example

```r
# trace test
denmark_ca <- ca.jo(denmark_ts, ecdet = "const", type = "trace", spec = "transitory", season = 4)
summary(denmark_ca)
```

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: trace statistic , without linear trend and constant in cointegration
##
## Eigenvalues (lambda):
## [1]  4.331654e-01  1.775836e-01  1.127905e-01  4.341130e-02 -7.250439e-16
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 3 |  2.35  7.52  9.24 12.97
## r <= 2 |  8.69 17.85 19.96 24.60
## r <= 1 | 19.06 32.00 34.91 41.07
## r = 0  | 49.14 49.65 53.12 60.16
```

Cointegration Test - Johansen's Methodology, An Example

```r
# max eigenvalue test
denmark_ca <- ca.jo(denmark_ts, ecdet = "const", type = "eigen", spec = "transitory", season = 4)
summary(denmark_ca)
```

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegrat
##
## Eigenvalues (lambda):
## [1]  4.331654e-01  1.775836e-01  1.127905e-01  4.341130e-02 -7.250439e-16
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 3 |  2.35  7.52  9.24 12.97
## r <= 2 |  6.34 13.75 15.67 20.20
## r <= 1 | 10.36 19.77 22.00 26.81
## r = 0  | 30.09 25.56 28.14 33.24
```

Cointegration Test - Johansen's Methodology, An Example

- trace test suggests that $\boldsymbol{y}_t$ are not cointegrated, we can't reject $H_0 : \text{rank}(\mathbf{\Pi}) = 0$
- max eigenvalue test however suggests that $\boldsymbol{y}_t$ are cointegrated with one cointegrating relationship, we can first reject $H_0 : \text{rank}(\mathbf{\Pi}) = 0$ and afterwards can't reject $H_0 : \text{rank}(\mathbf{\Pi}) = 1$

Cointegration Test - Johansen's Methodology, An Example

- function cajorols estimates the VEC model, given cointegration rank r

```
denmark_vec <- cajorls(denmark_ca, r = 1)
denmark_vec
```

```
## $rlm
##
## Call:
## lm(formula = substitute(form1), data = data.mat)
##
## Coefficients:
##           LRM.d      LRY.d      IBO.d      IDE.d
## ect1      -0.212955   0.115022   0.023177   0.029411
## sd1       -0.057653  -0.026826  -0.000400  -0.004830
## sd2       -0.016305   0.007842   0.007622  -0.001178
## sd3       -0.040859  -0.013083   0.004627  -0.002885
## LRM.dl1    0.262771   0.602668   0.057349   0.061340
## LRY.dl1   -0.144254  -0.142828   0.144224   0.017741
## IBO.dl1   -0.040115  -0.290609   0.310660   0.264939
## IDE.dl1   -0.670698  -0.182561   0.203769   0.212009
##
##
## $beta
##                ect1
## LRM.l1     1.000000
## LRY.l1    -1.032949
## IBO.l1     5.206919
## IDE.l1    -4.215879
## constant -6.059932
```

- more detailed output with standard errors, t statistics and p-values can be obtained using `summary(denmark_vec$rlm)`

Cointegration Test - Johansen's Methodology, An Example

- cointegrating vector is estimated as $\beta = (1, -1.03, 5.21, -4.22, -6.06)'$
- adjustment parameters are estimated as $\alpha = (-0.213, 0.115, 0.023, 0.029)'$ which is consistent with a stable error correcting mechanism
- note that we have included a constant in the cointegrating relationship

Deterministic Terms in VEC Model

five possible specifications of deterministic terms $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ in VEC model

1. $\boldsymbol{\mu}_t = \mathbf{0}$ (no constant)

$$\Delta \boldsymbol{y}_t = \boldsymbol{\alpha}\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

$y_{i,t}$ are $I(1)$ with no drift, cointegrating relationships $\boldsymbol{\beta}' \boldsymbol{y}_t$ have zero mean

2. $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 = \boldsymbol{\alpha}\boldsymbol{\delta}_0$ (restricted constant)

$$\Delta \boldsymbol{y}_t = \boldsymbol{\alpha}(\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \boldsymbol{\delta}_0) + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

$y_{i,t}$ are $I(1)$ with no drift, cointegrating relationships have non zero mean

3. $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$ (unrestricted constant)

$$\Delta \boldsymbol{y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

$y_{i,t}$ are $I(1)$ with drift, cointegrating relationships may have non zero mean

4. $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\delta}_1 t$ (restricted trend)

$$\Delta \boldsymbol{y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}(\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \boldsymbol{\delta}_1 t) + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

$y_{i,t}$ are $I(1)$ with drift, cointegrating relationships $\boldsymbol{\beta}' \boldsymbol{y}_t$ have linear trend

5. $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ (unrestricted trend)

$$\Delta \boldsymbol{y}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t + \boldsymbol{\alpha}\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{y}_{t-1} + \ldots + \boldsymbol{\Gamma}_{p-1} \Delta \boldsymbol{y}_{t-p+1} + \varepsilon_t$$

$y_{i,t}$ are $I(1)$ and have a drift and a quadratic trend, $\boldsymbol{\beta}' \boldsymbol{y}_t$ have linear trend

Example: Bivariate VEC with Deterministic Components

- let $\boldsymbol{y}_t = (y_{1,t}, y_{2,t})'$ follow a VEC model $\Delta \boldsymbol{y}_t = \boldsymbol{\mu}_t + \boldsymbol{\Pi} \boldsymbol{y}_{t-1} + \varepsilon_t$ with

$$\boldsymbol{\Pi} = \boldsymbol{\alpha}\beta' = \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix} \begin{bmatrix} 1 & -1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.2 \\ 0.2 & -0.2 \end{bmatrix}$$

  so that the adjustment parameters are $\alpha_1 = -0.2$, $\alpha_2 = 0.2$ and cointegrating relationship is $\beta' \boldsymbol{y}_t = y_{1,t} - y_{2,t}$

- consider now the following specifications of deterministic components
  - case 1: if $\boldsymbol{\mu}_t = 0$, there is no drift, and $E(\beta' \boldsymbol{y}_t) = 0$
  - case 2: if

$$\boldsymbol{\mu}_t = \begin{bmatrix} 0.4 \\ -0.4 \end{bmatrix}$$

    so that $\boldsymbol{\mu}_t = \boldsymbol{\alpha}\delta_0$ with $\delta_0 = -2$, there is no drift, and $E(\beta' \boldsymbol{y}_t) = -\delta_0 = 2$
  - case 3: if

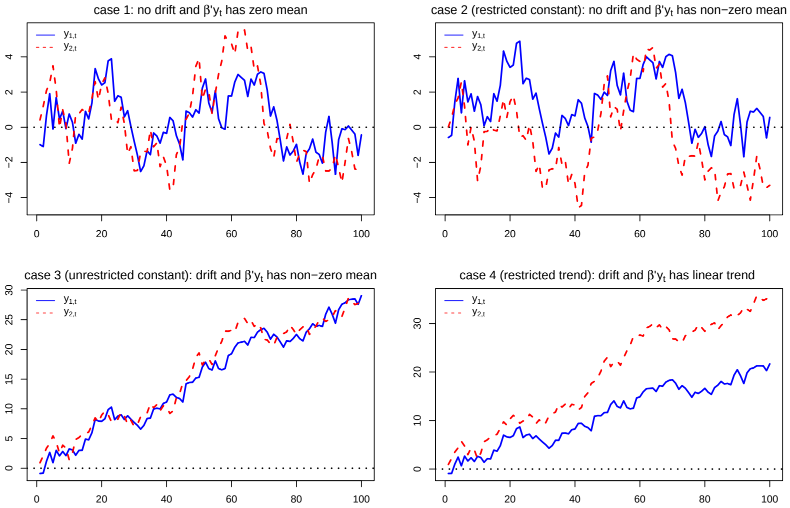
$$\boldsymbol{\mu}_t = \begin{bmatrix} 0.1 \\ 1.5 \end{bmatrix}$$

    then $y_{1,t}$ and $y_{2,t}$ have a drift and $E(\beta' \boldsymbol{y}_t) \neq 0$
  - case 4: if

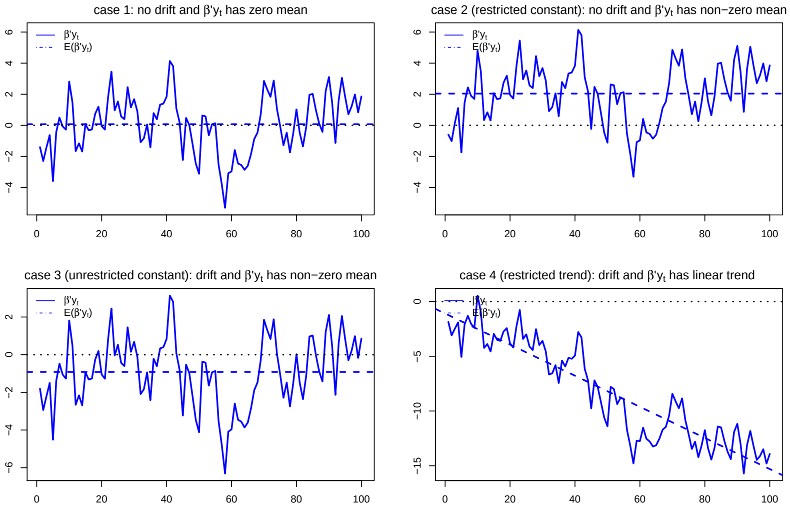
$$\boldsymbol{\mu}_t = \begin{bmatrix} 0.1 \\ 0.4 \end{bmatrix} + \begin{bmatrix} -0.3 \\ 0.3 \end{bmatrix} t$$

    so that $\mu_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\delta_1 t$ with $\delta_1 = 1.5$, there is drift, and also linear trend in cointegrating relationship, $E(\beta' \boldsymbol{y}_t) = -1.5t$

Example: Bivariate VEC with Deterministic Components



Example: Bivariate VEC with Deterministic Components



Specifying Lag Length and Deterministic Components

- lag length of VAR(p) can be determined using AIC, BIC, HQ
- function `ca.jo` from the `urca` package - use option `ecdet` to implement
  - case 2, restricted constant: `ecdet="const"`
  - case 3, unrestricted constant: `ecdet="none"`
  - case 4, restricted trend: `ecdet="trend"`
- as a rough rule of thumb
  - when *all* time series in $\boldsymbol{y}_t$ are non-trending like interest rates, exchange rates, inflation rate, unemployment rate, various growth rates, we use case 2
  - when one or more time series in $\boldsymbol{y}_t$ are trending, e.g. asset prices, macroeconomic aggregates like GDP, consumption, exports, industrial production, employment, national debt, M2 money stock, we start with case 4 or case 3, and can test whether we can impose restriction implied by case 2

Hypothesis Testing - Restrictions on Cointegrating Relationship

- one advantage of Johansen's approach is that it allows to easily test restrictions on $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$
- we can also test the specification of deterministic components:
  e.g. restricted constant in cointegrating relationship vs. presence of an unrestricted drift term
- main idea: if restrictions imposed are consistent with data and thus not binding, the number of cointegrating vectors stays same and rank($\mathbf{\Pi}$) stays the same

Hypothesis Testing - Restrictions on Cointegrating Relationship

- a test with H_0: restricted constant (case 2) against H_A : drift (case 3) is implemented using `lttest`

```
lttest(denmark_ca, r = 1)
```

```
## LR-test for no linear trend
##
## H0: H*2(r<=1)
## H1: H2(r<=1)
##
## Test statistic is distributed as chi-square
## with 3 degress of freedom
##         test statistic p-value
## LR test           1.98    0.58
```

- the results of the test above justify the specification used on previous slides where `ecdet="const"`

Hypothesis Testing - Restrictions on Cointegrating Relationship

- recall that in the Denmark money demand example where $\boldsymbol{y}_t = (\log(M2_t/P_t), \log Y_t, i_t^b, i_t^d)'$, the cointegrating relationship is

$$\beta_1 \log(M2_t/P_t) + \beta_2 \log Y_t + \beta_3 i_t^b + \beta_4 i_t^d + \beta_5$$

  and the cointegrating vector was estimated as $\boldsymbol{\beta} = (1, -1.03, 5.21, -4.22, -6.06)'$, so that β_2 is close to -1
- consider money demand $\frac{M}{p} = L(Y, i)$ and suppose that we wanted to test the hypothesis that $L(Y, i) = Y\tilde{L}(i)$ so that the velocity of money $v = pY/M$ is a function of interest rate i since $v = Y/(M/p) = 1/\tilde{L}(i)$
- money demand equation $\frac{M}{p} = Y\tilde{L}(i)$ where $\tilde{L}(i) = \gamma_0 e^{-\gamma_1 i}$ then implies a cointegrating relationship

$$\log \frac{M}{p} - \log Y + \gamma_1 i - \log \gamma_0$$

- this amounts to testing a restriction on the cointegration vector $\beta_2 = -\beta_1$, since with normalization $\beta_1 = 1$ we get $\beta_2 = -1$

Hypothesis Testing - Restrictions on Cointegrating Relationship

- to impose constraint $\boldsymbol{\beta}_2 = -\beta_1$ on the cointegrating vector let $\boldsymbol{\Psi} = (\psi_1, \psi_2, \psi_3, \psi_4)'$ and define matrix $\boldsymbol{H}$ as

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  so that the cointegrating vector is $\boldsymbol{\beta} = \boldsymbol{H}\boldsymbol{\Psi} = (\psi_1, -\psi_1, \psi_2, \psi_3, \psi_4)'$

- in general, to impose some linear constraints $\boldsymbol{R}'\beta = \mathbf{0}$ on cointegrating vectors, construct matrix $\boldsymbol{H}$ such that $\beta = \boldsymbol{H}\boldsymbol{\Psi}$ and use `blrtest` function from the `urca` package

Hypothesis Testing - Restrictions on Cointegrating Relationship

```r
# test for restricted cointegrating vector betta
rest_betta <- matrix(data = c(1,-1,0,0,0,
                              0,0,1,0,0,
                              0,0,0,1,0,
                              0,0,0,0,1),
                     nrow = 5, ncol = 4)
blrtest(denmark_ca, H = rest_betta, r = 1) %>% summary()
```

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Estimation and testing under linear restrictions on beta
##
## The VECM has been estimated subject to:
## beta=H*phi and/or alpha=A*psi
##
##      [,1] [,2] [,3] [,4]
## [1,]    1    0    0    0
## [2,]   -1    0    0    0
## [3,]    0    1    0    0
## [4,]    0    0    1    0
## [5,]    0    0    0    1
##
## Eigenvalues of restricted VAR (lambda):
## [1] 0.4327 0.1722 0.0436 0.0056
##
## The value of the likelihood ratio test statistic:
## 0.04 distributed as chi square with 1 df.
## The p-value of the test statistic is: 0.84
```

Hypothesis Testing - Restrictions on Adjustment Parameters

- restrictions on adjustment parameters $\boldsymbol{\alpha}$ can be implemented and tested in a similar way as restrictions on cointegrating vectors $\boldsymbol{\beta}$
- running `summary(denmark_vec$rlm)` shows that in the example with money demand for Denmark, α_1 looks significant but α_2, α_3, α_4 appear to be only marginally significant
- it thus makes sense to test the hypothesis $\alpha_2 = \alpha_3 = \alpha_4 = 0$
- to impose the above restriction let $\boldsymbol{\Psi} = \psi_1$ and define matrix $\boldsymbol{A}$ as

$$\boldsymbol{A} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

so that the adjustment vector is $\boldsymbol{\alpha} = \boldsymbol{A\Psi} = (\psi_1, 0, 0, 0)'$

- in general, to impose linear constraints $\boldsymbol{R}'\boldsymbol{\alpha} = \mathbf{0}$ on the adjustment parameters vectors, construct matrix $\boldsymbol{A}$ such that $\boldsymbol{\alpha} = \boldsymbol{A\Psi}$ and use `alrtest` function from the `urca` package

Hypothesis Testing - Restrictions on Adjustment Parameters

```r
# test for restricted adjustment parameters alpha
rest_alpha <- matrix(data = c(1,0,0,0), nrow = 4, ncol = 1)
alrtest(denmark_ca, A = rest_alpha, r = 1) %>% summary()
```

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Estimation and testing under linear restrictions on beta
##
## The VECM has been estimated subject to:
## beta=H*phi and/or alpha=A*psi
##
##      [,1]
## [1,]    1
## [2,]    0
## [3,]    0
## [4,]    0
##
## Eigenvalues of restricted VAR (lambda):
## [1] 0.3573 0.0000 0.0000 0.0000 0.0000
##
## The value of the likelihood ratio test statistic:
## 6.66 distributed as chi square with 3 df.
## The p-value of the test statistic is: 0.08
```

Hypothesis Testing - Restrictions on Cointegrating Relationship

- it is possible to test restrictions on α and β jointly using ablrtest

```
ablrtest(denmark_ca, A = rest_alpha, H = rest_betta, r = 1) %>% summary()
```

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Estimation and testing under linear restrictions on alpha and beta
##
## The VECM has been estimated subject to:
## beta=H*phi and/or alpha=A*psi
##
##      [,1] [,2] [,3] [,4]
## [1,]    1    0    0    0
## [2,]   -1    0    0    0
## [3,]    0    1    0    0
## [4,]    0    0    1    0
## [5,]    0    0    0    1
##
##
##      [,1]
## [1,]    1
## [2,]    0
## [3,]    0
## [4,]    0
##
## Eigenvalues of restricted VAR (lambda):
## [1] 0.3564 0.0000 0.0000 0.0000
##
## The value of the likelihood ratio test statistic:
## 6.73 distributed as chi square with 3 df.
## The p-value of the test statistic is: 0.08
```

Hypothesis Testing - Restrictions on Cointegrating Relationship

- to obtain a VEC with restrictions imposed on β use output of `blrtest` as input in function `cajorols`

```
denmark_ca_rest <- blrtest(denmark_ca, H = rest_betta, r = 1)
denmark_vec_rest <- cajorls(denmark_ca_rest, r = 1)
denmark_vec_rest
```

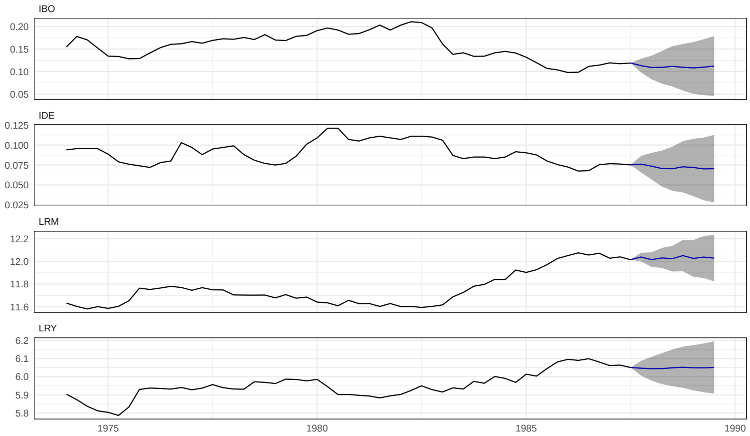
```
## $rlm
##
## Call:
## lm(formula = substitute(form1), data = data.mat)
##
## Coefficients:
##           LRM.d       LRY.d       IBO.d       IDE.d
## ect1      -0.2119917   0.1075103   0.0226379   0.0296896
## sd1       -0.0574493  -0.0265878  -0.0004005  -0.0048785
## sd2       -0.0164126   0.0076694   0.0076193  -0.0011495
## sd3       -0.0407508  -0.0130722   0.0046191  -0.0029036
## LRM.dl1    0.2556063   0.5999246   0.0577177   0.0627184
## LRY.dl1   -0.1379518  -0.1459504   0.1435555   0.0168536
## IBO.dl1   -0.0311275  -0.2721596   0.3111305   0.2623262
## IDE.dl1   -0.6646265  -0.1995151   0.2022586   0.2119758
##
##
## $beta
##                ect1
## LRM.l1     1.000000
## LRY.l1    -1.000000
## IBO.l1     5.300435
## IDE.l1    -4.290432
## constant -6.264457
```

Forecasting using a VEC model

- to construct forecasts, IRFs and FEVDs, we need to first transform the estimated VEC model in differences into a VAR in levels

```
denmark_var <- vec2var(denmark_ca, r = 1)
```

```
denmark_var_f <- predict(denmark_var, n.ahead = 8)
autoplot(denmark_var_f) + facet_wrap(~variable, ncol = 1, scales = "free_y")
```



Application: Pairs Trading

- cointegration and error correction model are used in the **pairs trading** strategy
- arbitrage pricing theory - if two stocks have similar characteristics, their prices must be more or less the same
- pairs trading involves selling the higher priced stock and buying the lower priced stock with the hope that the mispricing will correct itself in the future
- this strategy has been used on Wall Street for more than twenty years

Application: Pairs Trading

- consider two stocks with log prices $p_{i,t} = \log P_{i,t}$ for $i = 1, 2$ that follow random walk $p_{i,t} = p_{i,t-1} + r_{i,t}$ where $r_{i,t}$ are the serially uncorrelated log returns
- if the two stocks have similar risk factors, $p_{1,t}$ and $p_{2,t}$ will be driven by a common stochastic trend and thus cointegrated
- linear combination $w_t = p_{1,t} - \beta p_{2,t}$ will thus be $I(0)$ for some parameter β
- the stationary series w_t is referred to as the spread between the two log stock prices
- the two price series will follow error correction model

$$\begin{bmatrix} \Delta p_{1,t} \\ \Delta p_{2,t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \left[p_{1,t-1} - \beta p_{2,t-1} - \mu \right] + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}$$

- α_1 and α_2 should have opposite signs, indicating reversion to the equilibrium

Application: Pairs Trading

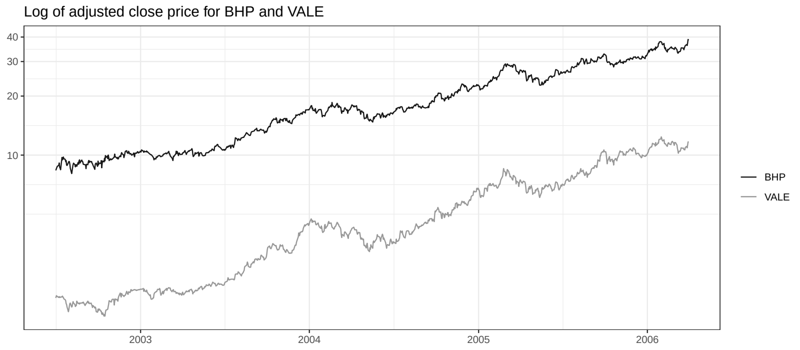
- since spread w_t is $I(0)$ it is mean reverting
- trade are carried out when $w_t = p_{1,t} - \beta p_{2,t}$ deviates substantially from its mean μ
- one possible trading strategy
  - buy a share of stock 1 and short β shares of stock 2 at time t if $w_t = \mu - \Delta$
  - unwind the position at time $t+i$ if $w_{t+i} = \mu + \Delta$
- here Δ is chosen such that $2\Delta > \eta$, where η is the costs of carrying out the two trades
- net profit is $2\Delta - \eta$
- a modified trading strategy: if $\Delta > \eta$ it is possible to unwind the position at time $t+i'$ if $w_{t+i'} = \mu$ which shortens the holding period of the portfolio

Application: Pairs Trading

```
library(tidyverse)
library(timetk)
library(vars)
library(urca)

theme_set(theme_bw())

# stock price data for Billiton Ltd. of Australia (BHP) and Vale S.A. of Brazil (VALE),
# two multinational companies in natural resources industry that face similar risk factors
# this data can be downloaded from http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/
webpage <- "http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/"

y_tbl <-
    inner_join(read_delim(file = str_c(webpage, "d-bhp0206.txt"), delim = " "),
               read_delim(file = str_c(webpage, "d-vale0206.txt"), delim = " "),
               by = c("Mon", "day", "year"), suffix = c("_BHP", "_VALE")) %>%
    gather(variable, value, -c("Mon", "day", "year")) %>%
    filter(str_sub(variable, 1, 8) == "adjclose") %>%
    mutate(date = (year*10000 + Mon*100 + day) %>% as.character() %>% as.Date("%Y%m%d" ),
           variable = str_sub(variable, 10, -1),
           logvalue = log(value))

# convert into zoo
y_zoo <-
    y_tbl  %>%
    dplyr::select(date, variable, logvalue) %>%
    spread(variable, logvalue) %>%
    tk_zoo(select = -date, date_var = date)
```

Application: Pairs Trading

```r
# time series plot - log of adjusted close price for BHP and VALE
y_tbl %>%
    ggplot(aes(x = date, y = value, col = variable)) +
        geom_line() +
        scale_y_log10(breaks = c(0,10,20,30,40)) +
        scale_color_manual(values = c("gray10","gray60")) +
        labs(x = "", y = "", col = "", title = "Log of adjusted close price for BHP and VALE")
```



Application: Pairs Trading

```r
# determine number of lags to be included in cointegration test and in VEC model
y_var_ic <- VARselect(y_zoo, type = "const")
nlags <- y_var_ic$selection["AIC(n)"]

# perform trace and maximum eigenvalue cointegration tests
y_ca <- ca.jo(y_zoo, ecdet = "const", type = "trace", K = nlags, spec = "transitory")
summary(y_ca)
y_ca <- ca.jo(y_zoo, ecdet = "const", type = "eigen", K = nlags, spec = "transitory")
summary(y_ca)
```

Application: Pairs Trading

```
##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: trace statistic , without linear trend and constant in cointegration
##
## Eigenvalues (lambda):
## [1] 4.148282e-02 8.206470e-03 2.825535e-18
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 1 |  7.78  7.52  9.24 12.97
## r = 0  | 47.77 17.85 19.96 24.60

##
## ######################
## # Johansen-Procedure #
## ######################
##
## Test type: maximal eigenvalue statistic (lambda max) , without linear trend and constant in cointegrat:
##
## Eigenvalues (lambda):
## [1] 4.148282e-02 8.206470e-03 2.825535e-18
##
## Values of teststatistic and critical values of test:
##
##           test 10pct  5pct  1pct
## r <= 1 |  7.78  7.52  9.24 12.97
## r = 0  | 40.00 13.75 15.67 20.20
```

Application: Pairs Trading

```r
# estimate VEC model
y_vec <- cajorls(y_ca, r = 1)
y_vec
```

```
## $rlm
##
## Call:
## lm(formula = substitute(form1), data = data.mat)
##
## Coefficients:
##           BHP.d     VALE.d
## ect1      -0.06731   0.02546
## BHP.dl1   -0.10949   0.06169
## VALE.dl1   0.07067   0.04768
##
##
## $beta
##               ect1
## BHP.l1    1.000000
## VALE.l1  -0.717704
## constant -1.828460
```

so the estimated VEC model takes form

$$\begin{bmatrix}\Delta p_{1,t}\\ \Delta p_{2,t}\end{bmatrix} = \begin{bmatrix}-0.067\\ 0.025\end{bmatrix}\left[p_{1,t-1} - 0.717 p_{2,t-1} - 1.828\right] + \begin{bmatrix}-0.109 & 0.071\\ 0.061 & 0.047\end{bmatrix}\begin{bmatrix}\Delta p_{1,t-1}\\ \Delta p_{2,t-1}\end{bmatrix} + \begin{bmatrix}\varepsilon_{1,t}\\ \varepsilon_{2,t}\end{bmatrix}$$

Application: Pairs Trading

- the spread is thus calculated as $w_t = p_{1,t} - \hat{\beta} p_{2,t} = p_{1,t} - 0.717 p_{2,t}$

```
# spread, its mean and standard deviation
w <- y_zoo %*% y_vec$beta[1:2]
mean(w)
```

```
## [1] 1.821159
```

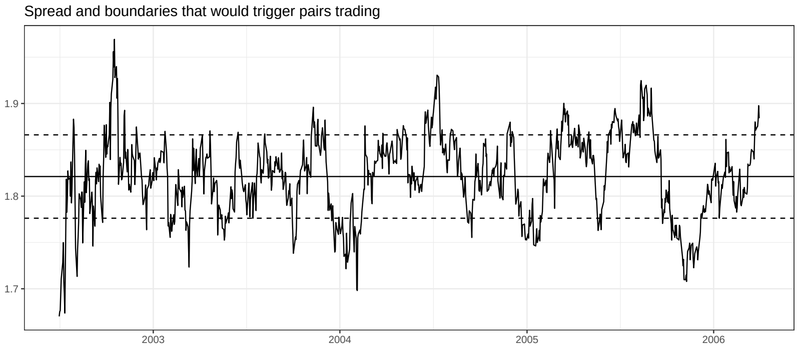
```
sd(w)
```

```
## [1] 0.04418623
```

- the mean spread is $\hat{\mu} = 1.821$
- the standard deviation is $\hat{\sigma} = 0.044$
- given that $\hat{\sigma}$ is quite large, it is possible to choose trading strategy by setting $\Delta = 0.045$ which yields log return for each pairs trading $2\Delta = 0.09$
- as shown in the figure on the next slide, w_t moves between $\hat{\mu} - 0.045$ and $\hat{\mu} + 0.045$ relatively often, so there are many pairs-trading opportunities

Application: Pairs Trading

```r
# plot spread and the boundaries that would trigger pairs trading
w %>%
    tk_tbl(rename_index = "date") %>%
    ggplot(aes(x = date, y = V1)) +
        geom_line() +
        geom_hline(yintercept = mean(w), linetype = "solid") +
        geom_hline(yintercept = mean(w) + 0.045, linetype = "dashed") +
        geom_hline(yintercept = mean(w) - 0.045, linetype = "dashed") +
        labs(x= "", y = "", title = "Spread and boundaries that would trigger pairs trading")
```



Application: Pairs Trading

- note that this illustrative example is based on in-sample analysis
- a realistic demonstration would require to assess the out-of-sample performance
- identifying cointegrated pairs of stocks that share similar risk factors may by quite challenging
- main issue: if a lot of traders exploit a particular pairs trading strategy, the stock may cease to be cointegrated

Summary VAR vs VEC

- if variables $\boldsymbol{y}_t$ are $I(0)$ we don't difference data and estimate VAR in levels
- if variables $\boldsymbol{y}_t$ are $I(1)$ we first test them for cointegration
  - if they are cointegrated we estimate a VEC model
  - if they are not cointegrated we difference the data and estimate a VAR model on first differences Δy_t